AstraZeneca

Applicant: **AstraZeneca AB**
S-151 85 Södertälje
Sweden

Title: **COMPOSITIONS AND METHODS UTILIZING SEQUENCES FOR CONTROLLING NUCLEIC ACID EXPRESSION IN YEAST**

Reference: F2215 (3526.82543)

Inventors: BELFIELD, Graham
OAKLEY, Caroline

BAK
2215 WO

## BACKGROUND OF THE INVENTION

The controlled production in yeast of an enormous variety of useful proteins or polypeptides can be achieved using recombinant DNA technology. Yeast cells can be transformed with yeast expression vectors, which contain homologous or heterologous nucleic acid molecules encoding polypeptides (coding sequences). The yeast cells can then produce large quantities of the useful proteins or polypeptides in yeast cell culture.

Expression of the nucleic acid molecule encoding a polypeptide by the yeast expression vector is initiated at a region known as the promoter, which is recognized by and bound by RNA polymerase. The RNA polymerase travels along the DNA, transcribing the information contained in the coding strand from its 5' to 3' end into messenger RNA, which is in turn translated into a polypeptide having the amino acid sequence for which the DNA codes. The present invention provides novel yeast promoters useful for, *inter alia*, controlling the expression of homologous and heterologous nucleic acid sequences encoding proteins and polypeptides in yeast cells.

## SUMMARY OF THE INVENTION

It is an object of the invention to provide novel yeast promoters, yeast expression vectors, and transformed yeast cells. It is a further object of the invention to provide a method for producing proteins and polypeptides in yeast cell culture.

In one embodiment of the invention a yeast promoter which comprises at least 17 contiguous nucleotides of an isolated and purified polynucleotide is provided. The promoter sequences are shown in SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, and SEQ ID NO:4.

The promoter is operative when operably linked to a nucleic acid molecule encoding a polypeptide.

As used herein, the term Apromoter@ refers to a nucleic acid sequence which is cable of initiating transcription of a nucleic acid molecule encoding a polypeptide (coding sequence); a Ayeast promoter@ is capable of initiating transcript of a coding sequence in yeast cells; and Apromoter activity@ refers to the level or amount of transcription initiation of a coding sequence, and encompasses any level above background (*i.e.*, the level or amount that occurs in the absence of a promoter; a background level, which is normally zero).

Another embodiment of the invention provides a yeast promoter which comprises an isolated and purified polynucleotide. The promoter sequences are shown in SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, and SEQ ID NO:4. The promoter is operative when operably linked to a nucleic acid molecule encoding a polypeptide.

Yet another embodiment of the invention provides a yeast promoter fragment which comprises at least 17 contiguous nucleotides of a polynucleotide. The polynucleotides are shown in SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, and SEQ ID NO:4. The fragment has promoter activity as determined by cloning the fragment into a yeast expression vector, wherein the fragment is operably linked to a reporter gene, transforming yeast cells with the yeast expression vector, growing the yeast cells in yeast cell culture under conditions favorable for expression of the reporter gene, and assaying the yeast culture for a reporter protein expressed by the reporter gene. The expression of the reporter gene indicates the fragment has promoter activity.

Still another embodiment of the invention provides a yeast expression vector comprising a yeast promoter. The promoter sequences are shown in SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, and SEQ ID NO:4. The promoter is operative when operably linked to a nucleic acid molecule encoding a polypeptide.

A further embodiment of the invention provides a yeast expression vector where activity of the promoter is controlled by varying the level of a non-fermentable carbon source, such as ethanol, in a medium of yeast cells in culture. The yeast cells are transformed with said yeast expression vector.

In yet another embodiment of the invention, a yeast expression vector comprising a yeast promoter which comprises at least 17 contiguous nucleotides of an isolated and purified polynucleotide is provided. The promoter sequences are shown in SEQ ID NO:1, SEQ ID NO:2, and SEQ ID NO:4. Promoter activity is controlled by varying the level of a fermentable carbon source in a medium of yeast cells in culture, where the yeast cells are transformed with the yeast expression vector. The fermentable carbon source can be glucose.

Another embodiment of the invention provides a yeast expression vector comprising a yeast promoter. The yeast promoter comprises at least 17 contiguous nucleotides of an isolated and purified polynucleotide. The promoter sequences are shown in SEQ ID NO:1, SEQ ID NO:2, and SEQ ID NO:4. Promoter activity is controlled by varying the level of a fermentable carbon source and a non-fermentable carbon source, such as ethanol, in a medium of yeast cells in culture, where the yeast cells are transformed with the yeast

expression vector. The fermentable carbon source can be glucose. The non-fermentable carbon source can be ethanol.

Still another embodiment of the invention provides a yeast cell transformed with a yeast expression vector. The yeast expression vector comprises a yeast promoter. The promoter sequences are shown in SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, and SEQ ID NO:4. The promoter is operative when operably linked to a nucleic acid molecule encoding a polypeptide.

Yet another embodiment of the invention provides a method for producing a polypeptide. A yeast expression vector is constructed where a polynucleotide encoding the polypeptide is controlled by a yeast promoter. The yeast promoter comprises at least 17 contiguous nucleotides of an isolated and purified polynucleotide. The promoter sequences are shown in SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, and SEQ ID NO:4. The promoter is operative when operably linked to a nucleic acid molecule encoding a polypeptide. A culture of yeast cells is transformed with the yeast expression vector. The yeast cells are maintained in culture so that the polypeptide is expressed. The polypeptide is then recovered.

Still another embodiment of the invention provides a method for producing a polypeptide. A nucleic acid molecule encoding the polypeptide is cloned into an expression vector selected from the group consisting of pYLR110P+luc, pYMR251AP+luc, pYMR107P+luc, pZEO1P+luc, pYLR110P, pYMR251AP, pYMR107P, and pZEO1P. The nucleotide acid molecule is operably linked to a promoter of the expression vector. A culture

of yeast cells is transformed with the yeast expression vector. The yeast cells are maintained in culture so that the polypeptide is expressed and the polypeptide is then recovered.

Another embodiment of the invention provides a method for producing a polypeptide. A yeast expression vector is constructed where a nucleic acid molecule encoding the polypeptide is controlled by a yeast promoter. The yeast promoter comprises at least 17 contiguous nucleotides of an isolated and purified polynucleotide. The promoter sequences are shown in SEQ ID NO:1, SEQ ID NO:2, and SEQ ID NO:4. Yeast cells are transformed with the yeast expression vector and are maintained in culture medium. The expression of the nucleic acid molecule encoding the polypeptide is controlled by varying the level of a fermentable carbon source, such as glucose, in the culture medium. The polypeptide is then recovered.

Still another embodiment of the invention provides a method for producing a polypeptide. A yeast expression vector is constructed where a nucleic acid molecule encoding the polypeptide is controlled by a yeast promoter. The yeast promoter comprises at least 17 contiguous nucleotides of an isolated and purified polynucleotide. The promoter sequences are shown in SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, and SEQ ID NO:4. The promoter is operative when operably linked to a nucleic acid molecule. A culture of yeast cells is transformed with the yeast expression vector. The yeast cells are maintained in culture medium and the expression of the nucleic acid molecule encoding the polypeptide is controlled by varying the level of a non-fermentable carbons source, such as ethanol, in the culture medium. The polypeptide is then recovered.

Another embodiment of the invention provides a method for producing a polypeptide. A yeast expression vector is constructed where a nucleic acid molecule encoding the polypeptide is controlled by a yeast promoter. The yeast promoter comprises at least 17 contiguous nucleotides of an isolated and purified polynucleotide. The promoter sequences are shown in SEQ ID NO:1, SEQ ID NO:2, and SEQ ID NO:4. A culture of yeast cells is transformed with the yeast expression vector. The yeast cells are maintained in culture medium and the expression of the nucleic acid encoding the polypeptide is controlled by varying the level of a fermentable carbon source, such as glucose, and a non-fermentable carbon source, such as ethanol, in the culture medium. The polypeptide is then recovered.

Yet another embodiment of the invention provides a method of identifying a promoter fragment with promoter activity by generating a fragment comprising at least 17 contiguous nucleotides of an isolated and purified polynucleotide. The polynucleotides are shown in SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, and SEQ ID NO:4. The fragment is cloned into a yeast expression vector, so that the fragment is operably linked to a reporter gene. Yeast cells are transformed with the yeast expression vector and grown in yeast cell culture under conditions favorable for expression of the reporter gene. The yeast culture is assayed for a reporter protein expressed by the reporter gene. Expression of the reporter gene indicates the fragment has promoter activity.

## BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 is a map of YEp13 expression vector.

Figure 2 schematically illustrates construction of YLR110C and YMR251WA promoter constructs.

Figure 3 is a map of pPRB1P.

Figure 4 is a map of pPRB1P+luc.

Figure 5 is a map of pYLR110P+luc.

Figure 6 is a is a map of pYMR251AP+luc.

Figure 7 is a map of pYMR107P+luc.

Figure 8 is a map of pZEO1P+luc.

Figure 9 is a map pYLR110P.

Figure 10 is a map of pYMR251AP.

Figure 11 is a map of pYMR107P.

Figure 12 is a map of pZEO1P.

Figure 13 schematically illustrates the YLR110C promoter region.

Figure 14 schematically illustrates the YMR251WA promoter region.

Figure 15 schematically illustrates the YMR107W promoter region.

Figure 16 schematically illustrates the ZEO1 promoter region.

## DETAILED DESCRIPTION OF THE INVENTION

Novel yeast promoters whose activity can be controlled by a fermentable carbon source, such as glucose, or a non-fermentable carbon source, such as ethanol, or both have been identified. The yeast promoters are useful for, *inter alia*, the high level production of proteins or polypeptides in yeast cell culture.

**Yeast Promoters**

The isolated and purified promoter polynucleotides of the invention are shown in SEQ ID NO:1 (the YLR110C promoter), SEQ ID NO:2 (the YMR251WA promoter), SEQ ID NO:3 (the YMR107W promoter), and SEQ ID NO:4 (the ZEO1 promoter). Yeast promoters comprising as little as 17 nucleic acids have been determined to function as promoters. The yeast promoters of the invention comprise at least 17, 25, 50, 75, 100, 150, 200, 250, 300, 350, 400, 450, 500, 600 or 700 contiguous nucleic acids of an isolated and purified polynucleotide up to the maximum length provided in any one of the sequences presented herein, that is, SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, and SEQ ID NO:4.

Preferably, the promoter polynucleotides are isolated free of other components, such as proteins and lipids. The polynucleotides can be made by a cell and isolated or can be synthesized in the laboratory, for example, using an automatic synthesizer or an amplification method such as PCR.

Naturally occurring variants and artificial sequence variants (that is, those which do not occur in nature) of the promoters are included in the invention. Variants of the promoters and/or fragments thereof have, along their entire length, sequence identity of at least 90%,

and preferably greater than 95% as determined by the Smith-Waterman homology search algorithm as implemented in MPsrch™ program (University of Edinburgh) using an affine gap search with the following search parameters: gap open penalty: 12, gap extension penalty: 1.

Fragments of the full-length promoters are also functional as promoters. A promoter fragment of at least 17 contiguous nucleotides may occur at any position along the full-length promoter as shown in SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3 or SEQ ID NO:4. Accordingly, promoter activity of 17 or more contiguous nucleotides occurring anywhere along the full-length promoter can be analyzed. Fragments of 17, 25, 50, 75, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650 or 700, nucleotides of the promoters may be constructed by, for example, subjecting an isolated promoter to restriction endonucleases, to 5'- or 3'-deletion mutagenesis, to PCR, or to site specific deletion. A combination of these methods can also be used to generate fragments of a promoter.

The invention further embodies a hybrid promoter, *i.e.*, a promoter that comprises more than one promoter or more than one fragment of a promoter from which it was derived. The promoter fragments can be derived from more than one of the promoter sequences shown in SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3 and SEQ ID NO:4. The promoters and fragments can be constructed as described above, ligated together, and cloned into a yeast expression vector. Where a promoter comprises nucleotides from at least two polynucleotides selected from the group consisting of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3 and SEQ ID NO:4, at least 5, 6, 7, 8, 9, 10, 25, 50, 75, 100, 150, 200, 250, 300, 350,

400, 450, 500, 550, 600, or 650 contiguous nucleotides are derived from each of the polynucleotides to form a promoter of at least 17 nucleotides. Alternatively, each of the full-length promoters can be combined with another full-length promoter or with fragments of another promoter.

The yeast promoters, fragments of the promoters, and hybrid promoters are useful for controlling expression of a protein or polypeptide when the yeast promoter is operably linked to a nucleic acid molecule encoding the protein or polypeptide.

**Determination of Promoter Activity**

Promoters and fragments of promoters can be assayed for promoter activity by cloning a fragment of a promoter, or a full-length promoter, or a hybrid promoter into a yeast expression vector so that is operably linked to a reporter gene, *i.e.*, a coding sequence for a reporter protein. The yeast expression vector is transformed in yeast cells, which are grown in yeast cell culture under conditions favorable for expression of the reporter gene, for example, under conditions providing a fermentable and/or non-fermentable carbon source. Expression of the reporter gene, as determined by an assay for the amount of a reporter protein expressed by the reporter gene, indicates that the promoter has activity.

For example, to determine if a promoter has activity, *i.e.* is operative, expression of a reporter gene by a promoter of the invention may be compared to expression of the reporter gene by a reference promoter such as PBR1 (Cottingham *et al.* (1991) Eur J Biochem 196(2):431-8; Sleep *et al.* (1991) Biotechnology 9(2):183-7; Finnis *et al.* (1992) Yeast 8(1):57-60; Meldgaard *et al.*(1995) Glycoconj J 12(3):380-90; Bach *et al.* (1996) Receptors

and Channels 4(2):129-39. A promoter, a fragment of a promoter, or a hybrid promoter of the invention is operative if it expresses at least 25% of the amount of a reporter protein as the full-length PBR1 promoter in a medium containing a non-fermentable carbon source, or a fermentable carbon source, or both. Preferably, an operative promoter expresses at least 50%, 75%,, 100%, 200%, 300%, 400%, or more of the amount of a reporter protein as the full-length PBR1 reference promoter.

Assays for promoter activity are useful for identifying yeast promoters with high activity and the specific nucleotide sequences of the promoters that are necessary for promoter activity.

**Yeast Expression Vectors**

The yeast promoters of the invention, which comprise isolated and purified polynucleotides selected from the group consisting of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, and SEQ ID NO:4 or fragments thereof, can be used to construct yeast expression vectors.

Yeast expression vectors are any vectors capable of autonomous replication within a yeast host organism or capable of integrating into the yeast genome. Yeast expression vectors are useful for introducing foreign DNA into yeast cells. Typical yeast expression vectors include yeast integrative plasmids (YIp), yeast replicating plasmids (YRp), yeast expression plasmids (YXp), yeast centromere-containing plasmids (YCp), and yeast episomal plasmids (YEp). Preferably, a yeast expression vector can be selected and maintained in both yeast and *E. coli*.

Yeast expression vectors, typically plasmids, incorporate the yeast promoters of the invention to control expression of nucleic acid molecules encoding heterologous or homologous proteins or polypeptides. The nucleic acid molecules are operably linked to a promoter in the yeast expression vector. A wide range of heterologous eukaryotic and prokaryotic proteins or peptides may be expressed by the vectors of the invention.

Expression vectors incorporating the promoters can be constructed by inserting into a vector a nucleic acid molecule encoding a protein or polypeptide (coding sequence) which is to be expressed. The coding sequence can be inserted at a restriction site which is provided downstream of a translation start codon controlled by the promoter. The coding sequence must be inserted in the correct translational reading frame.

Alternatively, the polynucleotide can itself be provided with a translational start codon followed directly by a coding sequence. Where the promoter does not contain a translational start codon, a restriction site is provided so that the coding sequence can be inserted in the correct reading fame and so that its translational start codon is correctly positioned in relation to the promoter. The coding sequence can encode heterologous or homologous or eukaryotic or prokaryotic polypeptides or proteins. In a preferred embodiment the coding sequence encodes a fusion protein. The coding sequence may further comprise a signal sequence.

In addition to the promoters of the invention, other components can be added to the expression vectors of the invention. For example, yeast selective markers, such as *LEU2* or *TRP1*, which allow for selection of yeast cells that have been effectively transformed by the

vector can be added. A yeast replication origin, such as the replication origin of the 2-micron plasmid or the autonomous ARS replication segment can be added. Upstream activating sequences and transcription terminator sequences may be added. Further, at least a portion of a bacterial plasmid, such as found in YEp13, can be added to enable the yeast expression vector to be manipulated in an intermediate bacterial host system, such as *Escherichia coli*.

The expression vector may also comprise a reporter gene which encodes, for example, β-galactosidase or luciferase. The reporter gene can be under the control of a promoter of the invention. Where the reporter gene, *i.e.*, coding sequence, is linked to a gene encoding a desired protein, assaying the level of expression of the reporter protein can quickly and easily determine the level of expression of the desired protein.

The expression vectors of the invention can be used to direct the fermentable carbon source- and/or non-fermentable carbon source-induced high level expression of proteins or polypeptides in yeast. The promoters of the invention can be induced by the presence of a fermentable carbon source, such as glucose, or a non-fermentable carbon source, such as ethanol, or both. That is, the promoters have greater promoter activity in the presence of a fermentable carbon source, or a non-fermentable carbon source, or both than in the absence of a fermentable carbon source, or a non-fermentable carbon source, or both. Promoters YLR110C, as shown in SEQ ID NO:1; YMR251WA, as shown in SEQ ID NO:2; and ZEO1, as shown in SEQ ID NO:4, can be induced by a fermentable carbon source, such as glucose, or by a non-fermentable carbon source, such as ethanol, or by both. Promoter YMR107W, as shown in SEQ ID NO:3, can be induced by a non-fermentable carbon source, such as

ethanol. Thus, the amount of expression of a homologous or heterologous nucleic acid molecule encoding a protein operably linked to the promoters of the invention can be controlled by varying the amount of an available fermentable carbon source, such as glucose, or a non-fermentable carbon source, such as ethanol, or both.

**Transformed Yeast Cells**

Yeast cells can be transformed with the yeast expression vectors of the invention. Transformation can be accomplished by well known methods, including, but not limited to electroporation, calcium phosphate precipitation, and microinjection. The yeast expression vectors of the invention can be used to transform yeast cells, including, but not limited to *Saccharomyces cerevisiae, S. uvarum, S. carlsbergensis, Saccharomycopsis lipolytica, Schizosaccharomyces pombe*, and *Kluyveromyces lactis*.

Transformed yeast cells containing a yeast expression vector can be grown in an appropriate medium for the yeast. A fermentable or non-fermentable carbon source can be added to the yeast culture medium in order to control the activity of the promoter.

Methods of Production of Proteins

Yeast cells transformed with expression vectors comprising a promoter of the invention can be used to produce proteins and polypeptides. Under proper cell culture conditions, preferably in the presence of a fermentable or non-fermentable carbon source, or both, the promoters of the invention will control expression of a nucleic acid molecule encoding a polypeptide operably linked to the promoter.

The protein or polypeptide can be retained within the yeast cell. The yeast cells can be then harvested, lysed, and the protein obtained and substantially purified in accordance with conventional techniques. Such techniques include, but are not limited to chromatography, electrophoresis, extraction, and density gradient centrifugation.

In a preferred embodiment of the invention, the protein or polypeptide to be recovered will further comprise a signal peptide capable of transporting the protein or polypeptide through the membrane of a transformed yeast cell. The protein or polypeptide can be recovered from the culture medium by, for example, adsorption or precipitation.

Further, the proteins and polypeptides may be produced as a fusion protein, which includes not only the amino acid sequence of the desired protein, but also one or more additional proteins. Affinity purification protocols can be used to facilitate the isolation of fusion proteins. Typically, a ligand capable of binding with high specificity to an affinity matrix is chosen as the fusion partner for the desired protein. For example, fusion proteins made with glutathione-S-transferase can be selectively recovered on glutathione-agarose and IgG-Sepharose can be used to affinity purify fusion proteins containing staphylococcal protein A.

Preferably, the protein or polypeptide of interest can be separated from the remainder of the fusion protein. The fusion protein can be constructed so that a site for proteolytic or chemical cleavage is inserted between the protein of interest and the fusion partner. For example, sites for cleavage by collagenase, Factor Xa protease, thrombin, and enterokinase, have been inserted between the fusion partner and the protein of interest. The protein of

interest can be also cleaved from the remainder of the fusion protein by chemical cleavage by, for example, hydroxylamine, cyanogen bromide (CNBr), or N-chlorosuccinamide.

The following are provided for exemplification purposes only and are not intended to limit the scope of the invention described in broad terms above. All references cited in this disclosure are incorporated by reference.

## EXAMPLE 1

## Preparation of Yeast Samples

### *S. cerevisiae strain 11C*

This example describes the growth of haploid *Saccharomyces cerevisiae* strain 11C. It has the genotype: *ade2*-161, *trp1*-Δ63, *ura3*-52, *lys2*-801, *leu2*Δ1 &/or *leu2*-3 &/or *leu2*-112, *his3*Δ200 &/or *his4*-519. 11C was generated by crossing the strains YPH500 (Mat a *ura3*-52 *lys2*-801 *ade2*-161 *trp1*-Δ63 *his3*Δ200 *leu2*Δ1) (Sikorski and Hieter. (1989) A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in *Saccharomyces cerevisiae*. *Genetics* 122: 19-27) and AH22 (MATa *leu2*-3 *leu2*-112 *his4*-519) (Hinnen *et al.* (1978) Transformation of yeast. *Proc. Natl. Acad. Sci. USA* 75: 1929-1933).

Three sterile 500 ml conical flasks, each containing 100 ml sterile YPD broth (Sigma, Cat No. Y-1375) were inoculated with sterile 10μl loops of differing quantities of the *S. cerevisiae* strain 11C from a freshly streaked YPD plate (Sigma, Cat No. Y-1500), and grown in an orbital shaker at 30°C, 200 rpm, overnight. The growth of 11C in the three flasks was measured by absorbance at 600nm. One flask was deemed to be at the late

exponential growth phase (1.98 ODU ml at 600 nm), and this culture was used to inoculate (50ml o/n culture per flask) 2 identical 5L sterile conical flasks (labeled E and L), each containing 1L sterile YPD broth to a final concentration of ~0.1 ODU ml. Flasks E and L were grown in an orbital shaker at 30°C, 200 rpm. 10ml samples were collected at times indicated below (Table 1). The samples were treated as follows: their growth was determined (A600nm), the possibility of contamination was checked (using a light microscope), cells were harvested in a benchtop centrifuge (~2000xg for 5 minutes), and the supernatant removed and frozen at -20°C (samples labeled E0 - E3, and L0 - L5).

Table 1. *Growth of cultures E and L as measure by absorbance at 600 nm.*

| **Time Point** | **Time after inoculation (min)** | **Growth of flask E (ODU)** | **Growth of flask L (ODU)** |
|---|---|---|---|
| T0 | 0 | 0.099 | 0.099 |
| T1 | 310 | 0.37 | 0.36 |
| T2 | 410 | 0.71 | 0.72 |
| T3 | 455 | 0.97 | 0.92 |
| T4 | 775 | - | 3.64 |
| T5 | 1420 | - | 6.05 |

After 455 minutes, a time deemed to be late exponential growth phase in glucose, flask E (*i.e.* early) was harvested (~2000xg for 5 minutes), split into 50ml aliquots, and frozen at -80°C. After 1420 minutes, a time deemed to be growth on ethanol, flask L (*i.e.* late) was harvested (~2000xg for 5 minutes), split into 50ml aliquots, and frozen at -80°C.

***Determination of Glucose and Ethanol concentration***

Supernatant samples (E0 - E3, and L0 - L5) were defrosted, and their ethanol and glucose contents were measured using ethanol (Boehringer, Cat. No. 176290) and glucose (Boehringer, Cat. No. 176251) detection kits according to manufacturers instructions. The concentrations determined are shown below in Table 2.

Table 2. *Glucose and Ethanol concentrations in supernatants of cultures E and L at different time points.*

| Sample | Time after inoculation (min) | Glucose level in media (g $L^{-1}$) | Ethanol level in media (g $L^{-1}$) |
|---|---|---|---|
| E0 | 0 | 20.0 | 0.0 |
| E1 | 310 | 21.8 | 0.3 |
| E2 | 410 | 21.8 | 0.8 |
| E3 | 455 | 21.2 | 0.87 |
| L0 | 0 | 20.0 | 0.0 |
| L1 | 310 | 22.2 | 0.36 |
| L2 | 410 | 22.0 | 0.62 |
| L3 | 455 | 20.0 | 0.87 |
| L4 | 775 | 11.8 | 5.2 |
| L5 | 1420 | 0.0 | 11.8 |

It can seen in Table 2 that at the point of culture harvest for E (E3, 455 minutes), the cells were still utilizing glucose as a carbon source, while at the point of culture harvest for L (L5, 1420 minutes), glucose was exhausted, and the cells were utilizing ethanol as a carbon source. Calibration values used to calculate glucose concentrations are shown in Table 3. Calibration values used to calculate ethanol concentrations are shown in Table 4.

Table 3. *Glucose standards*

| GLUCOSE STANDARDS g/l | OD A340 |
|---|---|
| 0 | 0 |
| 0.2 | 0.246 |
| 0.4 | 0.461 |
| 0.6 | 0.726 |
| 0.8 | 0.967 |
| 1 | 1.227 |

Table 4. *Ethanol standards*

| ETHANOL STANDARDS g/L | OD A340 |
|---|---|
| 4.72 | 0.041 |
| 9.44 | 0.083 |
| 18.88 | 0.166 |
| 37.76 | 0.322 |
| 56.6 | 0.534 |
| 75.5 | 0.664 |
| 94.4 | 0.846 |

## EXAMPLE 2

### Analysis of RNA Levels From Yeast Dimorphic Growth Samples

### *Total RNA Isolation*

Total RNA was isolated from 300ml of culture using the hot phenol protocol. The frozen yeast pellets were resuspended in lysis buffer (4ml) (0.5ml Tris-CL (1M, pH 7.5), 1.0 ml EDTA (0.5 M), 2.5ml 10% SDS, and 46.0ml $ddH_2O$) and an equal volume of acid phenol

was added and vortexed. Following incubation at 65°C for one hour (with occasional vigorous vortexing) the mixture was placed on ice for 10 minutes then centrifuged (10 minutes). The aqueous layer was transferred to a fresh centrifuge tube and mixed with an equal volume of phenol at room temperature. The mixture was centrifuged and an equal volume of chloroform was mixed with the aqueous layer in a fresh centrifuge tube. Following centrifugation the aqueous layer was transferred to a fresh centrifuge tube and sodium acetate (to a final concentration of 0.3M) and two volumes of 100% ethanol was added to precipitate the RNA. The mixture was placed at -20°C for 30 minutes then centrifuged for 10 minutes to pellet the RNA. The RNA pellet was washed 2-3 times with 70% ethanol then allowed to dry at room temperature. The pellet was resuspended in ddH2O (200-500 μL). The RNA was quantitated by measuring OD 260-280. Yield of total RNA was ~4.5mg from each culture.

***Poly A+ RNA Purification***

Poly A+ RNA was purified from total RNA using Qiagen Oligotex mRNA Midi Kit (Qiagen, Cat. No. 70042). 2mg of total RNA was used as starting material and made up to a volume of 500μl with DEPC treated $H_2O$. To this 500μl buffer OBB (2x binding buffer) and 55μl oligotex suspension was added. The AOligotex mRNA Spin-Column Protocol@ from the kit protocol booklet was followed. The pelleted mRNA was washed in 200μl 75% ethanol, dried and resuspended in 10μl DEPC treated $H_2O$. Yield of Poly A+ RNA was ~ 8μg for each sample.

***cDNA Synthesis***

cDNA was synthesized using the protocol for GeneChip Expression Analysis Manual using reagents from Gibco BRL Life Technologies Superscript Choice System cat. No. 18090-019. For each sample 5µg Poly A+ RNA was added to 100pmol of T7-$(dT)_{24}$ primer (sequence: GGCCAGTGAATTGTAATACGACTCACTATAGGGAGGCGG-(T)24, HPLC purified) (SEQ ID NO:15) in a total of 8µl (made up to volume with DEPC treated $H_2O$). The reaction mixture was incubated for 10 minutes at 70°C in a Perkin Elmer PE9600 thermalcycler then put on ice. The following reagents were added to the reaction mixture: 4µl 5x first strand cDNA buffer; 2µl 0.1M DTT; and 1µl 10mM dNTP mix. The reaction mixture was mixed and incubated at 37°C for 2 minutes in a Perkin Elmer PE9600 thermocycler. 5µl SuperScript II reverse transcriptase was then added. The mixture was incubated at 37°C for 1 hour in a Perkin Elmer PE9600 thermocycler.

The first strand cDNA reaction was placed on ice and the following reagents added: 91µl DEPC treated $H_2O$; 30µl 5x second strand reaction buffer; 3µl 10mM dNTP mix; 1µl 10units/µl *E. coli* DNA ligase; 4µl 10units/µl *E. coli* DNA Polymerase I; and 1µl 2units/µl RNase H. The mixture was incubated at 16°C for 2 hours in a Perkin Elmer PE9600 thermalcycler. 2µl 5units/µl T4 DNA Polymerase was then added. The mixture was incubated for a further 5 minutes at 16°C in a Perkin Elmer PE9600 thermalcycler. 10µl 0.5M EDTA was then added.

The double stranded DNA was cleaned up by phenol extraction. The reaction product transferred to a 1.5ml eppendorf tube and 162µl Tris pH 8.0 saturated phenol was

added. The tube was mixed by vortexing, the tube was then centrifuged in a microfuge at 13,000rpm for 5 minutes. The top fraction was recovered and cDNA precipitated by addition of 60µl 7.5M ammonium acetate plus 400µl absolute ethanol. This was immediately centrifuged in a microfuge at 13,000rpm for 20 minutes. The supernatant fraction was discarded, the pellet was washed in 75% ethanol and then air-dried. The pellet was resuspended in 20µl DEPC treated $H_2O$

***Synthesis of Biotin-Labeled cRNA by In Vitro Transcription (IVT)***

Reagents from Ambion MEGAscript T7 kit, cat. No. 1334, were used for the synthesis of biotin-labeled cRNA by *in vitro* transcription (IVT). The NTP Labeling mix comprised 7.5mM ATP; 7.5mM GTP; 5.625mM UTP; 1.875mM Biotin-16-UTP (Enzo cat No. 42814); 5.625mM CTP; and 1.875mM Biotin-11-CTP (Enzo cat No. 42818). The IVT Labeling reaction comprised: 14.5µl NTP Labeling mix; 2µl 10x Ambion Transcription Buffer; 1.5µl Double strand cDNA (from above); and 2µl Ambion T7 Enzyme Mix.

The reaction mixture was incubated for 6 hours at 37°C in a Perkin Elmer PE9600 thermalcycler. The biotinylated cRNA was cleaned up using Qiagen RNeasy kit, cat No. 74103. The RNeasy kit protocol was followed exactly. RNA was eluted in 2 aliquots of 30µl DEPC treated $H_2O$. The RNA was precipitated by addition of 6µl 3M sodium acetate pH 5.5 plus 75µl absolute ethanol. The RNA was allowed to precipitate overnight at -20°C. Samples were centrifuged in a microfuge at 13,000rpm for 20 minutes to pellet the RNA. The supernatant fraction was discarded and the pellet was washed in 1ml of 75% ethanol

and then allowed to air dry. The pellet was then resuspended in 20μl DEPC treated $H_2O$. The yield of cRNA was ~ 40μg for each sample.

***cRNA Fragmentation***

11μg of cRNA was fragmented. 8μl of 5x Fragmentation buffer (200mM Tris-Acetate pH 8.1, 500mM potassium acetate, 150mM magnesium acetate) plus 11μg cRNA made up to 20μl with DEPC treated $H_2O$ was used. The reaction mixture was incubated 94°C for 35 minutes in a Perkin Elmer PE9600 thermal cycler.

***Hybridization to GeneChip Microarray***

The hybridization mix comprised: 20μl (11μg) of fragmented cRNA; 2.2μl of control oligo B2 (50pmol/μl) (sequence: 5=Biotin-GTCAAGATGCTACCGTTCAG 3= HPLC purified) (SEQ ID NO:16); 2.2μl Herring Sperm DNA (10mg/ml); 110μl 2x Buffer (2M NaCl, 20mM Tris pH 7.6, 0.01% Triton X-100); and 85.6μl DEPC treated $H_2O$. The hybridization mix heated to 95°C in a Techne hot block for 5 minutes, followed by incubation at 40°C for 5 minutes. The hybridization mix was clarified by centrifugation in microfuge at 13,000rpm for 5 minutes.

200μl of supernatant to added to the GeneChip cartridge (GeneChip cartridge was previously pre-wetted with 200μl 1x Buffer and incubated for 10 minutes at 40°C in the rotisserie box of a GeneChip hybridization oven 320 (cat No. 800127) at maximum rpm. The sample was hybridized to the microarray overnight at 40°C in a GeneChip hybridization oven in the rotisserie at maximum rpm.

***Washing and Staining of Probe Arrays***

The hybridization mix was recovered from the GeneChip cartridge and put back in the tube containing the remainder of the sample. 200μl 6x SSPE-T (6x SSPE plus 0.005% Triton X-100) was applied to the chip and pipetted in and out twice. This process was repeated twice more. Another 200μl 6x SSPE-T was applied to the cartridge and the cartridge was then incubated for 1 hour at 50°C at maximum rpm in the GeneChip hybridization oven. The 6x SSPE-T was removed and 200μl 0.5x SSPE-T was added to cartridge. The cartridge was incubated for 15 minutes at 50°C at maximum rpm in the GeneChip hybridization oven. The 0.5x SSPE-T was removed and the cartridge was re-filled with 200μl 6x SSPE-T.

The stain solution comprised: 190μl 6x SSPE-T; 10μl of 20mg/ml acetylated BSA; and 2μl 1mg/ml conjugated streptavidin:phycoerythrin (Molecular Probes cat. No. S-866). 200μl 6x SSPE-T was removed from the GeneChip cartridge and 200μl of stain solution added. The cartridge was incubated at ambient temperature in a GeneChip hybridization oven at maximum rpm in the rotisserie for 10 minutes. The stain solution was removed and the cartridge was washed by adding 200μl 6x SSPE-T and pipetting this in and out of the cartridge twice. This process was repeated six times. The cartridges were then completely filled with 6x SSPE-T and any bubbles removed. Hybridization, washing and staining was repeated using the same hybridization mixes until both samples had been hybridized to each of the four yeast chip sub-set arrays.

***Data collection***

Data was collected by scanning the hybridized chips on a Hewlett-Packard GeneArray scanner. A Ahalo@ effect (appearance of stain non-specifically across the array image) was seen on one of the scanned images: yeast growing in glucose rich media, sub-set C array. Scanning of this array was aborted after one scan and the chip was washed twice with 200µl 6x SSPE-T and then re-filled as before. This array was then re-scanned three times and the data collected was the average of these three scans. All other arrays were scanned four times without problems and the data collected was the average of the four scans.

**EXAMPLE 3**

Isolation of promoters and construction of expression vectors.

*PCR amplification of promoter regions from genomic DNA*

Based on the *Saccharomyces cerevisiae* genomic sequence in the GenEMBL nucleotide database oligonucleotide primers were designed to amplify the genomic sequence 5= to the following ORFs: YLR110C (Johnston *et al.* (1997) Nature 1997 May 29;387(6632 Suppl):87-90), YMR251WA (common name HOR7) (Bowman *et al.* (1997) Nature May 29;387(6632 Suppl):90-3), YMR107W (Bowman *et al.* (1997) Nature May 29;387(6632 Suppl):90-3), and YOL109W (common name ZEO1) (Dujon *et al.* (1997) Nature May 29;387(6632 Suppl):98-102). The region amplified was the non-coding region separating the selected ORF and the next predicted *Saccharomyces cerevisiae* ORF in the 5= direction, with a minimum length of 500bp.

***Sequence of oligonucleotide primers used to amplify promoter DNA***

HindIII, NheI and NdeI cloning sites underlined.

| | | |
|---|---|---|
| YLR110C-F | ATGC<u>AAGCTT</u>CGCGGCCGCCGTCTGATTTCCGTTT | SEQ ID NO:5 |
| YLR110C-R | CCAGGCCG<u>CATATG</u>TCATATAGTGTTTAAG | SEQ ID NO:6 |
| YMR251WA-F | AGCT<u>AAGCTT</u>CGCGGCCGCCTTTCGATTAGCACGCAC | SEQ ID NO:7 |
| YMR251WA-R | AGATACCTT<u>CATATG</u>TTATTATTAGTC | SEQ ID NO:8 |
| YMR107W-F | AGCT<u>AAGCTT</u>CGCGGCCGCGCAGAAATGATGAAGG | SEQ ID NO:9 |
| YMR107W-R | ATCCATCC<u>CATATG</u>TGATATCTCGATTAG | SEQ ID NO:10 |
| ZEO1-F | AGCT<u>AAGCTT</u>CGCGGCCGCGGAGGTCTGCTTCACG | SEQ ID NO:11 |
| ZEO1-R | TACGATCG<u>CATATG</u>TAATTGATATAAACG | SEQ ID NO:12 |

PCR reactions were set up for each primer pair as follows: For YMR251WA and ZEO1 90µl of Reddy-Load PCR (1.1X) mix, 3.5mM $MgCl_2$ (Advanced Biotechnologies, cat.no. AB-0628); 2µl of forward primer (100µM); 2µl of reverse primer (100µM); 1µl of *S. cerevisiae* genomic DNA (Promega G310A, lot 8347702, 276µg/ml); and 5µl of $H_2O$ were combined.

For YLR110C and YMR107W 90µl of Reddy-Load PCR (1.1X) mix, 1.5mM $MgCl_2$ (Advanced Biotechnologies, cat.no. AB-0575); 2µl of forward primer (100µM); 2µl of reverse primer (100µM); 1µl of *S. cerevisiae* genomic DNA (Promega G310A, lot 8347702, 276µg/ml); and 5µl of $H_2O$ were combined.

The thermocycling was carried out as follows: For the YMR251WA promoter: 94°C for 5 minutes followed by 30 cycles of : 94°C for 30 seconds, 60°C for 30 seconds, 72°C for 1 minute; followed by 72°C for 5 minutes. The reaction mixtures were then held at 4°C. For the YMR107W and ZEO1 promoters: 94°C for 5 minutes followed by 30 cycles of : 94°C for 30 seconds, 45°C for 30 seconds, 72°C for 1 minute; followed by 72°C for 5 minutes.

The reaction mixtures were then held at 4°C. For the YLR110C promoter: 94°C for 5 minutes followed by 30 cycles of : 94°C for 30 seconds, 50°C for 30 seconds, 72°C for 1 minute; followed by 72°C for 5 minutes. The reaction mixtures were then held at 4°C.

The PCR solutions were loaded onto an LMP gel and the bands were purified using Wizard PCR Preps (Promega, cat. no. A7170) according to protocol, eluted in 50µl, ethanol precipitated, and resuspended in 20µl. A map of the YLR110C promoter region is shown in Figure 13 and SEQ ID NO:29. A map of the YMR251WA promoter region is shown in Figure 14 and SEQ ID NO:30. A map of the YMR107W promoter region is shown in Figure 15 and SEQ ID NO:31. A map of the ZEO1 promoter region is shown in Figure 16 and SEQ ID NO:32.

***Cloning promoter regions into a yeast vector containing the luciferase gene***

The PCR products representing the regions upstream of the YLR110C and YMR251WA ORFs were cloned into the suitably digested YEp13-based multicopy yeast expression vector pPRB1P+luc. A map of YEp13 is shown in Figure 1. The Accession number for YEp13 is U03498. A map of pPRB1P is shown in Figure 2. The sequence of pPRB1P is shown in SEQ ID NO:27. A map of pPRB1P+luc is shown in Figure 3 and the sequence is shown in SEQ ID NO:28. The PRB1 promoter was removed from the vector by digesting with the restriction enzymes HindIII and NdeI. The digested backbone was then ligated with a HindIII / NdeI digested PCR product. See Figure 4.

The PCR products described below, and maxi-prepped pPRB1P+luc were digested as follows. 60 µl of pPRBP1+luc (328µg/ml), 10 µl of Hind III (Life Technologies, cat.no.

15207-012, 10 units/μl), 10 μl NdeI (Amersham, cat.no. E0216Y, 20 units/μl), 10 μl NEBuffer 2 (NEB, cat.no. 007-2), and 10 μl of $H_2O$. 14 μl YLR110C, 2 μl of Hind III (Life Technologies, cat.no. 15207-012, 10 units/μl), 2 μl Nde I (Amersham, cat.no. E0216Y, 20 units/μl), and 2 μl NEBuffer 2 (NEB, cat.no. 007-2). 14 μl YMR251WA, 2 μl of Hind III (Life Technologies, cat.no. 15207-012, 10 units/μl), 2 μl Nde I (Amersham, cat.no. E0216Y, 20 units/μl), and 2 μl NEBuffer 2 (NEB, cat.no. 007-2). The solutions were allowed to react at 37°C, for 4 hours.

The double digested pPRB1P+luc backbone was purified on an LMP gel using Wizard PCR preps (Promega, cat. no. A7170), and then ethanol precipitated. The remaining digestion products were also ethanol precipitated. The pPBR1P+luc digests were resuspended in 60μl of $H_2O$ and the PCR product digests were resuspended in 20μl.

Ligation reactions were then carried out between each promoter region and the digested pPRBP1+luc at 16°C overnight. The PCR products representing the regions upstream of the following ORFs; YMR107W and ZEO1, were prepared, restricted, and ligated essentially as described above, however BCL restriction buffer B and different amounts of PCR product/volumes were used.

***<u>Transformation of ligation products into E. coli</u>***

The products of the ligations described above were transformed into *E. coli* (Invitrogen=s One-Shot TOP10 Competent cells, cat.no. C4040-10) according to manufacturers protocol. In each case 5μl of the ligation product was added to the cell

suspension. The total final cell suspension was plated out onto L-amp plates and incubated overnight at 37°C.

Colonies were picked from the plates and PCR screened using the PCR primers used to amplify the promoters originally. Two positive colonies from each ligation were grown in 5ml overnight cultures and their plasmids were purified (Promega Wizard Plus SV Mini-preps, cat. no. A1330). The eluted DNA was ethanol precipitated and resuspended in 20µl of water. Analytical restriction digests were carried out to confirm the presence of the correct promoter. Clones containing all four promoter constructs were obtained.

The new constructs were named as follows:

| | |
|---|---|
| pPRB1+luc backbone + YLR110C promoter = pYLR110P+luc | SEQ ID NO:19 |
| pPRB1+luc backbone + YMR251WA promoter = pYMR251AP+luc | SEQ ID NO:20 |
| pPRB1+luc backbone + YMR107W promoter = pYMR107P+luc | SEQ ID NO:21 |
| pPRB1+luc backbone + ZEO1 promoter = pZEO1P+luc | SEQ ID NO:22 |

Maps of pYLR110P+luc, pYMR251AP+luc, pYMR107P+luc, and pZEO1P+luc are shown in Figures 5, 6, 7, and 8, respectively. Plasmid DNA (pYLR110P+luc and pYMR251AP+luc) was prepared for transformation into yeast and sequencing using the QIAGEN Plasmid Maxi kit (Cat.no. 12162). The DNA concentrations of the maxi-preps (measured by absorbance at 260 nm) were: pYLR110P+luc 463µg/ml; pYMR251AP+luc 346 µg/ml; pYMR107P+luc ~300µg/ml; and pZEO1P+luc ~720 µg/ml. The remaining plasmids were transformed into yeast as Wizard Plus SV Mini-prep DNA, and maxi-prep DNA was obtained for sequencing using the Gibco BRL Concert Plasmid Maxi kit (Cat no.11452).

***Sequencing of promoter constructs***

DNA of each of the four promoter constructs were sequenced using the ABI PRISM BigDye Terminator Cycle Sequencing Kit (PE Applied Biosystems, part no. 4303153) was used to carry out the sequencing reactions. Each reaction contained 8μl of Reaction Mix and 1μl of 3.2 μM primer. The volumes of template DNA and $H_2O$ added are as follows: 1.1μl of pYLR110P+luc template and 9.9μl of water; 1.4μl of pYMR251AP+luc template and 9.6μl of water; 2.0-6.0μl of pYMR107P+luc template and 9.0-5.0μl of water; and 0.5-1.5μl of pZEO1P+luc template and 10.5-9.5μl of water.

The thermocycling protocol is described in the ABI protocol, the PCR products were ethanol precipitated by adding 3M NaOAc and absolute Ethanol, standing at room temperature for 15 minutes, centrifuging for 20 minutes and washing with 250μl of 70% ethanol. The precipitated DNA was resuspended in 3μl of loading dye and 2μl of each suspension was analyzed on an PE-ABI 377 automated sequencer.

The following promoter constructs pYLR110P+luc and pYMR251AP+luc were each sequenced using four primers:

YEp13 F2 : CCTCAATTGGATTAGTCTCA - SEQ ID NO:13- aligns to the YEp13 backbone, 290bp 5= of the Hind III site.

Luc R1 : CACCTCGATATGTGCATCTG - SEQ ID NO:14- aligns to the Luc ORF, 150bp 3= of the NdeI site.

Forward PCR primer: forward primer used to PCR clone promoter, *i.e.*, SEQ ID NO:5 and SEQ ID NO: 7.

Reverse PCR primer: reverse primer used to PCR clone promoter , *i.e.*, SEQ ID NO:6 and SEQ ID NO:8.

The remaining promoter constructs (pYMR107P+luc and pZEO1P+luc) were each sequenced using primers Yep13 F2 and Luc R1. Combining the data from all primers completely sequenced the promoter regions and spanned the cloning sites of the original vector.

***Deviations from published genomic sequences***

All sequences differ by a few base pairs around the ATG, this results from the creation of an NdeI site at the 3= end of the promoter. In addition, the following further alterations from published sequences were identified.

pYLR110P+luc: A substitution of a C for a T had taken place at a base pair 361 of the sequence.

pYMR107P+luc: In the initial construct (for which luciferase reporter data is described), a cloning artifact led to the junction between the promoter region and the LUC ORF in pYMR107W+luc to have the sequence: CATATATG (where ATG is the luciferase translational start site). This sequence was modified by site directed mutagenesis to create the sequence CATATG , which generates a novel NdeI site at the promoter/luciferase junction. Subsequent luciferase expression analysis confirmed that expression from the NdeI site modified pYMR107P+luc construct did not differ significantly from the original construct, therefore the sequence of the corrected CATATG construct is included herein.

***Other Modifications***

pYMR107P+luc: Cloning artifacts created an additional HindIII site and linker to the 5= (*i.e.* outside ) of the pYMR107P+luc and promoters:

Instead of:

**hindIII** **NotI** **promoter 5=**
AAGCTT-CGCGGCCGCG-NNNNNNNN SEQ ID NO:17

The sequence is:

**hindIII** **hindIII** **NotI** **promoter 5=**
AAGCTT-AGCT-AAGCTT-CGCGGCCGCG-NNNNNNNN SEQ ID NO:18.

**EXAMPLE 4**

***Luciferase assays of promoter activity***

Transformation of *S. cerevisiae* with promoter constructs.

*S. cerevisiae* strain 11C was transformed with five promoter constructs. This strain carries six metabolic markers, Ade, Trp, Ura, Lys, Leu and His. It has the genotype: *ade2*-161, *trp1*-D63, *ura3*-52, *lys2*-801, *leu2*D1 &/or *leu2*-3 &/or *leu2*-112, *his*D200 &/or *his*D200. 11C was generated by crossing the strains YPH500 (Mat a *ura3*-52 *lys2*-801 *ade2*-161 *trp1*-D63 *his*D200 *leu2*D1) and AH22 (MATa *leu2*-3 *leu2*-112 *his4*-519 *can1*.

11C cells were streaked from a glycerol stock onto a YPD plate and grown at 30°C for two days. The cells were transformed with the five plasmids, pYLR110P+luc, pYMR251AP+luc, pYMR107P+luc, & pZEO1P+luc and pPRB1P+luc to act as a control. The transformations were carried out using the Quick and Easy method (Gietz, R.D. and R.A. Woods, 1994, *Molecular Genetics of Yeast: Practical Approaches* pp. 121-134. 10ml of plasmid was added to the transformation mix in each case. The whole transformation mixes were plated out onto -Leu plates and incubated at 30°C for three days. Three individual colonies from each transformation plate were picked and used to inoculate 10ml YPD cultures. The 10ml cultures were incubated in an orbital shaker set to 200rpm and 30°C. Cells were harvested from the cultures at two points. First, at a point at which the OD of the culture was close to 1.0, at which time a 4ml sample was taken. Second, a 3ml sample was taken after an incubation time of 45 hours. The ODs and incubation time of each sample is shown in Table 5. For all harvested samples, the cells were immediately spun down at 3000rpm and 4°C, washed in 5ml of $dH_2O$, repelleted and frozen at -20°C.

Table 5

| Plasmid | Clone number | OD at time of harvesting first 4ml sample | Incubation time at harvesting of first sample (hours) | OD at time of harvesting second 3ml sample |
|---|---|---|---|---|
| PPRB1P | 7 | 0.98 | 24.5 | 4.80 |
| +luc | 8 | 0.68 | 28 | 5.56 |
| | 9 | 1.15 | 28 | 5.66 |
| PYLR110P | 8 | 1.12 | 28 | 5.50 |
| +luc | 9 | 0.46 | 28 | 4.38 |
| | 10 | 1.16 | 24.5 | 5.51 |
| PYMR251AP | 8 | 1.20 | 24.5 | 4.99 |
| +luc | 9 | 1.05 | 27 | 4.71 |
| | 10 | 1.15 | 27 | 5.18 |
| PYMR107P | 1 | 1.06 | 27 | 5.47 |
| +luc | 2 | 0.49 | 28.5 | 4.54 |
| | 3 | 0.97 | 25.5 | 5.58 |
| PZEO1P | 1 | 1.02 | 28.5 | 4.84 |
| +luc | 2 | 0.62 | 28.5 | 4.97 |
| | 3 | 0.42 | 28.5 | 4.31 |

***Analysis of luciferase activity***

All of the samples were analyzed for luciferase activity, using the LucLite Luciferase Reporter Gene Assay Kit (Packard, cat.no 6016911). The cells were prepared by resuspending in PBS and diluting to a final concentration of $6 \times 10^6$ cells/ml. 100ml of each cell suspension was pipetted into wells in duplicate on two 96 well plates, so that each well contained $6 \times 10^5$ cells. The plates were incubated at 30°C for 10 minutes. 100ml of a 1 in 2 dilution of reconstituted substrate was added to each well, and the plate was further incubated at room temperature for 10 minutes. The luminescence was then

measured using the Packard TopCount. The luminescence readings obtained after 0.03min are shown below in counts per second (CPS) in Table 6.

Table 6.

| Plasmid | Clone number | First sample | | | | Second sample | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Readings (CPS) | | Average | Average | Readings (CPS) | | Average | Average |
| PPRB1P +luc | 7 | 35890 | 35690 | 35790 | 34898 | 20322 | 20975 | 20648 | 19867 |
| | 8 | 25498 | 25276 | 25387 | 24495 | 52997 | 51778 | 52388 | 51607 |
| | 9 | 24137 | 27797 | 25967 | 25075 | 49192 | 46971 | 48081 | 47300 |
| PYLR110P +luc | 8 | 52354 | 53618 | 52986 | 52094 | 41789 | 38904 | 40346 | 39565 |
| | 9 | 105299 | 99776 | 102537 | 101645 | 85562 | 84468 | 85015 | 84234 |
| | 10 | 107531 | 109226 | 108379 | 107486 | 22507 | 22436 | 22471 | 21690 |
| PYMR251AP +luc | 8 | 71993 | 69797 | 70895 | 70003 | 40869 | 40202 | 40536 | 39755 |
| | 9 | 98853 | 98389 | 98621 | 97729 | 51159 | 49828 | 50493 | 49712 |
| | 10 | 83210 | 87546 | 85378 | 84485 | 70091 | 74576 | 72334 | 71553 |
| PYMR107P +luc | 1 | 9046 | 8650 | 8848 | 6790 | 29413 | 28505 | 28959 | 28124 |
| | 2 | 3996 | 4009 | 4002 | 1945 | 24391 | 23915 | 24153 | 23318 |
| | 3 | 3018 | 3236 | 3127 | 1069 | 23866 | 23408 | 23637 | 22802 |
| PZEO1P +luc | 1 | 64137 | 63162 | 63649 | 61592 | 47469 | 45769 | 46619 | 45784 |
| | 2 | 19579 | 18329 | 18954 | 16897 | 44910 | 42982 | 43946 | 43111 |
| | 3 | 87572 | 90317 | 88944 | 86887 | 142414 | 142262 | 142338 | 141503 |

The results are summarized in Table 7.

Table 7.

| Promoter | mRNA levels | Luciferase Expression Glucose | Luciferase Expression Ethanol |
|---|---|---|---|
| PRB1 | Ethanol Induced | 1.00 | 1.00 |
| YLR110C | Highly Ethanol and Glucose Induced | 3.03 | 1.22 |
| YMR251WA | Highly Ethanol and Glucose Induced | 2.92 | 1.35 |
| YMR107W | Ethanol Induced | 0.21 | 0.95 |
| ZEO1 | Very Highly Ethanol | 3.62 | 2.89 |

| | | | |
|---|---|---|---|
| | and Glucose Induced | | |

Three promoters give higher levels of expression than PRB1 at both ODs, these are: YLR110C, YMR251WA, and ZEO1. The promoter showing the greatest fold induction is YMR107W.

***Creating vectors with promoters but without the luciferase gene***

Based on the analysis of luciferase expression four further promoter constructs have been made. The lack the luciferase gene and can be used to clone nucleic acid molecules encoding polypeptides of interest downstream of the promoters such that they drive expression of the nucleic molecules of interest. The sequences of these four plasmids are named: G1: pYLR110P (SEQ ID NO:23) (map at Figure 9); G2: pYMR251AP (SEQ ID NO:24) (map at Figure 10); G3 pYMR107P (SEQ ID NO:25) (map at Figure 11); and G4: pZEO1P (SEQ ID NO:26) (map at Figure 12). These were constructed by digesting pPRB1P (SEQ ID NO:27) with HindIII and NdeI to obtain the vector. The promoter+luc construct was digested with HindIII and NdeI to obtain the promoter fragment. The vector and promoter DNA was purified from LMP agarose using PCRpreps. The vector and promoter was ligated and used to transform *E. coli*. Correct recombinants were screened for.

**EXAMPLE 5**

***Isolation of Active Promoter Fragments***

Operative fragments of the YLR110C, YMR251WA, YMR107W and ZEO1 promoters can be generated using restriction endonucleases, 5' or 3' deletion mutagenesis, PCR, site specific deletion, or a combination thereof. For example, purified pYLR110P+luc,

pYMR251AP+luc, pYMR107P+luc or pZEO1P+luc plasmids, as generated in Example 3, can be subjected to restriction endonucleases to generate fragments of the YLR110C, YMR251WA, YMR107W or ZEO1 promoters. Restriction endonuclease sites, preferably unique restriction endonuclease sites, within the promoter sequences shown in SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, and SEQ ID NO:4 can be identified that generate fragments of the promoter upon restriction endonuclease digestion. Such fragments are preferably, 17, 25, 50, 75, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650 or 700 nucleotides in length.

The fragments generated by restriction endonuclease digestion of the promoters shown in SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, or SEQ ID NO:4 can be separated by agarose gel electrophoresis. The agarose gel band corresponding to the desired promoter fragment can be cut out of the agarose gel. The fragment can be isolated and purified from the agarose gel by, for example, electroelution or kits such as QIAquick™ gel extraction kit or QIAEX7 II Gel Extraction System (Qiagen Cat. No. 28704 and 20021).

The purified promoter fragment can be ligated into the isolated and purified HindIII, NdeI, double-digested pPRBP1+luc backbone such that the promoter fragment is operably linked to a luciferase gene and transformed into *E. coli*, as described in Example 3. The new expression vector comprising a fragment of YLR110C, YMR251WA, YMR107W, or ZEO1 promoter region can be isolated and purified from *E. coli*, sequenced, and transformed into yeast as described in Example 3.

To analyze promoter activity, luciferase assays as described in Example 4, can be conducted using *S. cerevisiae* cultures that have been transformed with the expression vector comprising a fragment of the YLR110C, YMR251WA, YMR107W, or ZEO1 promoter operably linked to a luciferase gene and *S. cerevisiae* cultures that have been transformed with pPRB1P+luc. The *S. cerevisiae* cultures are grown in medium containing a non-fermentable carbon source, such as ethanol, or a fermentable carbon source, such as glucose, or both. Cells are obtained from the cultures and analyzed for luciferase activity as described in Example 4.

A promoter fragment is operative if it expresses at least 75% of the luciferase activity as the PRB1 promoter. Preferably, an operative promoter fragment expresses at least 100%, 200%, 300%, 400%, or more of the luciferase activity as the PRB1 promoter.

**BRIEF DESCRIPTION OF THE SEQUENCES**

SEQ ID NO:1 Polynucleotide sequence of promoter YLR110C

SEQ ID NO:2 Polynucleotide sequence of promoter YMR251WA

SEQ ID NO:3 Polynucleotide sequence of promoter YMR107W

SEQ ID NO:4 Polynucleotide sequence of promoter ZEO1

SEQ ID NO:5 Forward PCR primer for YLR110C

SEQ ID NO:6 Reverse PCR primer for YLR110C

SEQ ID NO:7 Forward PCR primer for YMR251WA

SEQ ID NO:8 Reverse PCR primer for YMR251WA

SEQ ID NO: 9 Forward PCR primer for YMR107W

SEQ ID NO:10 Reverse PCR primer for YMR107W

SEQ ID NO:11 Forward PCR primer for ZEO1

SEQ ID NO:12 Reverse PCR primer for ZEO1

SEQ ID NO:13: Yep13 Forward PCR primer

SEQ ID NO:14: Luc RI Forward PCR primer

SEQ ID NO:15 Primer used in cDNA sequencing

SEQ ID NO:16 Control oligonucleotide used in GeneChip Microarray assay

SEQ ID NO:17 Original pYMR107P+luc sequence

SEQ ID NO:18 Modified pYMR107P+luc sequence

SEQ ID NO:19 Nucleotide sequence of pYLR110P+luc

SEQ ID NO:20 Nucleotide sequence of pYMR251AP+luc

SEQ ID NO:21 Nucleotide sequence of pYMR107P+luc

SEQ ID NO:22 Nucleotide sequence of pZEO1P+luc

SEQ ID NO:23 Nucleotide sequence of pYLR110P

SEQ ID NO:24 Nucleotide sequence of pYMR251AP

SEQ ID NO:25 Nucleotide sequence of pYMR107P

SEQ ID NO:26 Nucleotide sequence of pZEO1P

SEQ ID NO:27 Nucleotide sequence of pPRB1P

SEQ ID NO:28 Nucleotide sequence of pPRB1P+luc

SEQ ID NO:29 YLR110C promoter region

SEQ ID NO:30 YMR251WA promoter region

SEQ ID NO:31 YMR107W promoter region

SEQ ID NO:32 ZEO1 promoter region